B. Amendments to the Claims

1 1.   (Amended) A method of cooperatively load-balancing a cluster of server
2 computer systems for servicing client requests issued from a plurality of client
3 computer systems, wherein said client computer systems are mutually independent
4 in issuing client requests to said cluster of server computer systems and wherein
5 said cluster of server computer systems interoperate with said plurality client
6 computer systems to establish a cooperative load-balancing of said cluster of
7 server computer systems, said method comprising the steps of:
8   a) selecting, by a client computer system, a target server computer system
9 from said cluster of server computer systems to service a particular client request
10 using available accumulated selection basis data, wherein said available
11 accumulated selection basis data is accumulated from said cluster of server
12 computer systems;
13   b) evaluating, by said target server computer system, said particular client
14 request to responsively provide instance selection basis data dynamically
15 dependent on the configuration of said target server computer and said particular
16 client request; and
17   c) incorporating, by said client computer system, said instance selection
18 basis data into said available accumulated selection basis data to affect a
19 subsequent selection of said target computer system with respect to a subsequent
20 instance of said particular client request.

1 2.   (Original) The method of Claim 1 wherein said instance selection basis
2 data includes a representation of a dynamically determined performance level of
3 said target server computer system and wherein said available accumulated
4 selection basis data incorporates said instance selection basis data with
5 identifications of said target server computer and said particular client request.

1 3.   (Original) The method of Claim 2 wherein said instance selection basis
2 data includes a representation of a policy evaluation of said particular client
3 request relative to said target server computer system.

1 4.   (Original) The method of Claim 1 wherein said instance selection basis
2 data includes a load value and a selection weighting value, wherein said load

3   value represents a dynamically determined performance level of said target server
4   computer system and said selection weighting value represents a policy evaluation
5   of said particular client request relative to said target server computer system and
6   wherein said available accumulated selection basis data incorporates said
7   instance selection basis data with identifications of said target server computer
8   and said particular client request.

1   5.   (Original) The method of Claim 4 wherein said step of selecting selects said
2   target server computer system based on predetermined selection criteria including
3   the relative values of said load value and said selection weighting value with
4   respect to said particular client request as recorded in said available accumulated
5   selection basis data.

1   6.   (Original) The method of Claim 5 wherein said instance selection basis
2   data provides for a rejection of said particular client request and wherein said step
3   of selecting includes selecting an alternate server computer system from said
4   cluster of server computer systems as said target server system to service said
5   particular client request based on said available accumulated selection basis data.

1   7.   (Previously Presented) A method of load-balancing a cluster of server
2   computer systems in the cooperative providing of a network service to host
3   computers operating mutually independent of one another, said method
4   comprising the steps of:
5   a) selecting, independently by each of a plurality of host computers, server
6   computers within a computer cluster to which to issue respective service requests;
7   b) responding, by a corresponding one of said plurality of host computers,
8   to a rejection of a predetermined service request by selecting a different server
9   computer to which to issue said predetermined service request;
10   c) receiving, in regard to said respective service requests by respective ones
11   of said plurality of host computers, load and weight information from respective
12   said server computers, wherein load and weight information is dynamically
13   generated by respective said server computers; and
14   d) evaluating, by each of said plurality of host computers, respective load
15   and weight information received with respect to said server computers of said

16   computer cluster as a basis for a subsequent performance of said step of
17   selecting.

1   8.   (Previously Presented) The method of Claim 7 further comprising a step of
2   determining said weight information by each of said server computers with respect
3   to each service request received, said weight information being determined from
4   a predefined policy association between a received service request and the identity
5   of the one of said server computers that receives the service request.

1   9.   (Previously Presented) The method of Claim 8 further comprising a step of
2   distributing initial information by said cluster of server computers to said host
3   computers, said initial information providing selection lists of said server
4   computers to said host computers.

1   10.   (Original) The method of Claim 9 wherein said load information is
2   representative of a plurality of load factors including network loading and
3   processor loading.

1   11.   (Previously Presented) The method of Claim 10 wherein said load
2   information is representative of processing a current set of service requests
3   including a plurality of processor functions.

1   12.   (Original) The method of Claim 11 wherein said load information includes
2   one or more load values representing processing functions internal to a server
3   computer.

1   13.   (Previously Presented) A server cluster operated to provide a load-balanced
2   network service, said server cluster comprising:
3   a) a plurality of server computers individually responsive to service requests
4   to perform corresponding processing services, wherein said server computers are
5   operative to initially respond to said service requests to provide load and weight
6   values, wherein said load and weight values represent a current operating load
7   and a policy-based priority level of a respective server computer relative to a
8   particular service request; and

9   b) a host computer system operative to autonomously issue said service
10   requests respectively to said plurality of server computers, said host computer
11   system further operative to select a target server computer from said plurality of
12   server computers to receive an instance of said particular service request based
13   on said load and weight values.

1   14.   (Previously Presented) The server cluster of Claim 13 wherein said host
2   computer is operative to collect said load and weight values from said plurality of
3   server computers in connection with respective service requests issued to said
4   plurality of server computers and wherein the selection of said target server
5   computer is based on relative temporal ages of said load and weight values.

1   15.   (Original) The server cluster of Claim 14 wherein each of said plurality of
2   server computers include a policy data set store that provides for the storage of
3   a distinct server configuration and wherein said load and weight values are
4   dynamically determined by said plurality of server computers in response to said
5   service requests based on said distinct server configurations of said plurality of
6   server computers.

1   16.   (Original) The server cluster of Claim 15 wherein said distinct server
2   configurations include the distinct identities of said plurality of server computers.

1   17.   (Original) The server cluster of Claim 16 wherein said distinct server
2   configurations include distinct policy data relative to said service requests, wherein
3   said host computer system is operative to collect, relative to respective said service
4   requests, and provide attribute data to said plurality of server computers, and
5   wherein said server computers evaluate said attribute data in conjunction with said
6   distinct policy data to determine said weight values.

1   18.   (Original) The server cluster of Claim 17 wherein said plurality of server
2   computers implement a security processing service, wherein said host computer
3   system is operative to selectively route network transported data through said
4   server computers dependent on said service requests as evaluated by said plurality
5   of server computers.

1   19.   (Previously Presented) The server cluster of Claim 18 wherein said host
2   computer is operative to initiate respective data transfer transactions for each of
3   said service requests, wherein a default routing of each said data transfer
4   transaction initially provides for the transfer of corresponding ones of said service
5   requests to respective ones of said plurality of server computers, and wherein said
6   respective ones of said plurality of server computers determine whether the
7   subsequent routing of network data within said respective data transfer
8   transactions includes routing said network data within said respective data transfer
9   transactions through said plurality of server computers.

1   20.   (Previously Presented) A computer system providing, in support of a
2   plurality of client computer systems, a network service through a scalable cluster
3   of server computer systems, said system comprising:
4   a) a first plurality of server computers coupled to provide a defined service,
5   wherein a server computer of said first plurality provides a response, including
6   dynamically determined load and weight information, in acknowledgment of a
7   predetermined service request issued to said server computer system, said
8   response selectively indicating nonacceptance of said predetermined service
9   request; and
10   b) a second plurality of client computers coupleable through a network to
11   said first plurality, said second plurality including a client computer system having
12   an identification list of said first plurality, said client computer system being
13   operative to autonomously select a first server computer system from said
14   identification list to which to issue said predetermined service request, wherein
15   said client computer system is reactive to said response, on indicated
16   nonacceptance of said predetermined service request, to autonomously select a
17   second server computer system from said identification list to which to issue said
18   predetermined service request, and wherein said client computer system is
19   responsive to said load and weight information of said response in subsequently
20   autonomously selecting said first and second server computer systems.

1   21.   (Cancelled).

1   22.   (Previously Presented) The computer system of Claim 20 wherein said first
2   plurality of server computer systems include respective policy engines and wherein

3   said weight information reflects an association between a server computer policy

4   role and said predetermined service request.

1   23.   (Original) The computer system of Claim 22 wherein said predetermined

2   service request includes predetermined client process attribute information and

3   wherein said respective policy engines are responsive to said predetermined client

4   process attribute information in determining said server computer policy role

5   relative to said predetermined service request.

1   24.   (Original) The computer system of Claim 23 wherein said load information

2   includes a value representing network and server processor performance.

1   25.   (Previously Presented) A method of dynamically managing the distribution

2   of client requests from a plurality of mutually independent client computer systems

3   to a plurality of server computer systems providing a network service, each of said

4   server computer systems being discretely configured to respond to client requests,

5   said method comprising the steps of:

6   a) processing client requests to select for a particular client request a

7   particular server computer system of said plurality of server computer systems to

8   service said particular client request, wherein the selection of said particular server

9   computer system is dependent on the evaluation of accumulated selection

10   qualification information;

11   b) forwarding said particular client request to said particular server

12   computer system; and

13   c) receiving from said particular server computer system with respect to said

14   particular client request instance selection qualification information discretely

15   determined by said particular server computer system dynamically with respect to

16   said particular client request, wherein said instance selection qualification

17   information including a load value reflective of the current performance capability

18   of said particular server computer system and a weight value reflective of the

19   anticipated performance capability of said particular server computer system with

20   respect to said particular client request, wherein said instance selection

21   qualification information is incorporated into said accumulated selection

22   qualification information.

1   26.   (Original) The method of Claim 25 wherein said processing step
2   dynamically evaluates said particular client request with respect to said
3   accumulated selection qualification information to identify said particular server
4   computer system as a best choice of said plurality of server computer systems for
5   selection.

1   27.   (Previously Presented) The method of Claim 26 further comprising a step
2   of evaluating by said particular server computer system, subject to the discrete
3   configuration of said particular server computer system, said particular client
4   request to provide said weight value as part of said instance selection qualification
5   information.

1   28.   (Cancelled).

1   29.   (Previously Presented) The method of Claim 27 wherein said weight value
2   part of said instance selection qualification information includes a relative
3   prioritization of said particular client request with respect to said particular server
4   computer system.

1   30.   (Original) The method of Claim 29 wherein said client requests are issued
2   with respect to client computer systems, wherein said particular client request
3   includes attributes descriptive of a particular client computer system that issued
4   said particular client request, and wherein said relative prioritization reflects the
5   evaluation of said attributes with respect to said particular server computer system.

1   31.   (Previously Presented) A method of distributing computational load over a
2   plurality of server systems provided to support execution of a data processing
3   service on behalf of a plurality of client systems, wherein the computational load
4   is generated in response to client requests issued through a plurality of client
5   processes, said method comprising the steps of:
6   a) first processing a particular client request to associate attribute data from
7   a respective client process of said plurality of client processes with said particular
8   client request;
9   b) selecting, for said particular client request, a particular target server
10   system from among said plurality of server systems by matching said particular

11   client request against accumulated selection information to identify said particular
12   target server system;
13   c) second processing said particular client request, including said attribute
14   data, by said particular target server system to dynamically generate instance
15   selection information including a load value for said particular target server
16   system and reflective of a combination of said particular client request and said
17   particular target server system and a relative weighting value reflective of the
18   combination of said particular client request and said particular target server
19   system; and
20   d) incorporating said instance selection information into said accumulated
21   selection information for subsequent use in said step of selecting, wherein said
22   step of selecting matches said particular client request, including said attribute
23   data, against corresponding data of said accumulated selection information to
24   choose said particular target server system based on a best corresponding
25   combination of relative weighting value and load value.

1   32.   (Cancelled).

1   33.   (Previously Presented) The method of Claim 31 wherein said step of
2   selecting includes a step of aging said accumulated selection information.

1   34.   (Original) The method of Claim 33 further comprising the steps of:
2   a) first providing, through a host process, said particular client request,
3   including attribute data, to said particular target server system; and
4   b) receiving by said host process, a particular target server response
5   including said instance selection information;
6   c) determining, by said host process from said particular target server
7   response, whether to select an alternate target server system;
8   d) reselecting, for said particular client request, a secondary target server
9   system from among said plurality of server systems by matching said particular
10   client request against said accumulated selection information, including said
11   instance selection information received from said particular target server response
12   to identify said secondary target server system; and
13   e) second providing, through said host process, said particular client
14   request, including attribute data, to said alternate target server system.

1   35.   (Original) The method of Claim 34 wherein said host process is executed

2   on a client computer system.

1   36.   (Original) The method of Claim 35 wherein said host process is executed

2   on a gateway computer system coupleable through a communications network

3   with a plurality of client computer systems.